Filtering and Aggregation Schemes for Delay Model Calibration

Andrew M. Churchill (churchil@umd.edu)

Kleoniki Vlachou (kvlachou@umd.edu)

David J. Lovell (lovell@umd.edu)

Department of Civil and Environmental Engineering & Institute for Systems Research
University of Maryland College Park
College Park, MD

Abstract—This paper describes some methods for filtering and aggregating delay data from individual flights. The purpose for these transformations is to make the delay data more consistent with the outputs from queuing models. The transformed data can then be used to make much more relevant, and successful, comparisons against such models. This enables better calibration of the models, and helps to reveal what fraction of the total delay in a system might be generated solely from the consideration of congestion resulting from competition amongst aircraft for scarce airspace and airport resources. The paper describes the transformations in detail, and demonstrates their theoretical validity through examples. Real data are modified according to these transformations and are then compared against a stochastic queuing model to show the efficacy of the technique.

Keywords-queuing models, airport delay, delay filtering

I. INTRODUCTION

Queuing models, either deterministic or stochastic, are commonly used to predict delay statistics in the National Airspace System (NAS). The need for estimating delay is great, especially for busy airports. These models are particularly useful for studying future conditions that might include changes from the demand and capacity profiles expected under current operations. In some cases, such models can also be used to predict the effects of important infrastructure or policy changes, such as the addition of a runway or changing separation standards.

Queuing models are designed to estimate that component of delay that is incurred by aircraft as a result of competition, with other aircraft, for a capacitated resource, such as a portion of the airspace or a runway. Real delay data at a destination airport represent a broader collection of influences, and might include, among other things, upstream delays that accrued in previous flight legs, delays caused by late arrival of the crew, and delays caused by a mechanical problem with an aircraft.

Queuing models expect, as their inputs, nominal arrival times of flights; i.e., those arrival times that would prevail if other influences did not create delays. The differences between the nominal (or "desired") arrival times and the actual arrival times are the statistics recorded as delays. Real demand data, in the form of flight schedules, do not represent this notion exactly, primarily because air carriers include in their estimates of arrival time some expectation of delays. When executed properly, this is a perfectly reasonable practice, because it maximizes the likelihood that the actual performance of a flight will match the customers' expectations for that flight. The problem, however, arises when an analyst tries to use the same data to populate queuing models, because their intent is to estimate those congestion-related delays themselves, rather than having them subsumed in the input data source. In this paper we offer a partial solution to this problem, although it is our belief that the general question of determining nominal (not padded for expected delay) arrival times for aircraft remains open.

All of the above constitute some of the reasons that the process necessary to facilitate proper comparisons of the outputs of queuing models with real data can be quite involved. Some form of comparison is essential, however, because the queuing models require calibration. It is also important to understand, once they are calibrated, that their delay predictions represent only a fraction of the total delay that might be expected when all of the other influences (which might be more difficult to model) are present. This proportion can be estimated as part of the overall calibration process and that is a valuable result in and of itself.

II. APPROACHES FOR MODIFYING DATA

The methods described in this paper for transforming individual flight data can be thought of as belonging to two classes of operations: filtering and aggregation. In the former case, we are attempting to subtract from real flight delay data the best estimates of delay components that are not directly attributable to congestion in the queuing sense. This step makes a direct comparison with delay data from queuing models much more valid.

Because the results from queuing models are most often shown in aggregate terms (e.g., the average delay incurred by all aircraft in the system during a particular time slice), the second step is to aggregate the real data, filtered accordingly, in a manner consistent with how queuing models tend to report their results. The need for this step is obvious, but its inclusion here is important because the paper illustrates how the

ISBN: 978-0-615-20720-9

aggregate delay statistics available in the most common aviation databases are *not* averaged in a manner that allows for direct comparison with model results. Any attempt to make such a comparison, therefore, without following steps such as those outlined in this paper, is very likely to lead to a poor match between model results and real data, leading to the possible (and likely erroneous) conclusion that the model is doing a bad job or that queuing delays are not a significant component of the overall delays incurred by aircraft.

A. Filtering Schemes

The basic inputs to a queuing model are demands and capacities. A straightforward (although, we will argue, incorrect) method to use real data to feed such a model would be to use the collection of scheduled arrival times at an airport as the demand and a record of the declared airport arrival rates (AARs) as the capacity. The outputs from the queuing model might include average delay per time period, and it might be tempting to compare these directly against an aggregate average delay statistic in a database such as the FAA Aviation System Performance Metrics (ASPM) database, partly because the name of the metric is very similar. Again, this paper offers evidence that a more refined method is better for these purposes.

The filtering mechanism encapsulates two basic processes, one for the input data for the model, and one for real delay data to which output data will be compared. In the input data, rather than using scheduled arrival times directly, we develop a scheme for predicting the nominal or "best" arrival time for each flight being considered. Since the queuing model only represents congestion effects at the single airport in question, data of similar scope must be used for output comparisons. We take individual flight delay data from a real database and subtract an estimate of upstream propagated delays that would not be accountable for in the queuing models. These processes are described in detail in Section III.

B. Aggregation Schemes

When looking at one of the readily available aviation performance databases such as ASPM, one can find aggregate delay statistics recorded on an hourly (or sometimes quarter hourly) basis. For example, one could find a report of the average delay at Atlanta Hartsfield-Jackson International Airport (ATL) between 4 pm and 5 pm on some day. It is not clear simply from the title of the field, however, what the domain of aggregation is. In fact, what happens, using the above example, is that for all flights that landed at ATL between 4 pm and 5 pm, their delays (relative to schedule) were computed, and then these were averaged over these flights. A flight scheduled to land at 3 pm but landing at 5 pm would be assigned two hours of delay, but both of those hours of delay would be aggregated in the time window 4-5 pm, when actually only one of them actually occurred during that window. In fact, given the possibility of upstream propagated delays, the actual delays might have occurred considerably earlier in the day. This is not a flaw in the reporting mechanism, however; the way that ASPM (and other) delays are aggregated is simply the easiest and least ambiguous way to record the ultimate differences between scheduled and actual arrival times.

The problem comes when trying to compare such data to the outputs of queuing models. In a deterministic queuing model, one can track the progress of individual aircraft, so it is possible to generate data that are consistent with this reporting mechanism. It is more common with such models, however, to use delay accounting practices taken directly from seminal sources on deterministic queuing (see for example [1]), where delays are accounted for *as they occur*, rather than after flights have landed. This difference can perhaps best be seen by graphical example; we call the mechanism used for reporting real data in places like ASPM "horizontal aggregation" and that typically used in deterministic queuing "vertical aggregation."

Importantly, stochastic queuing models (not simulations) frequently do not allow for the tracking of individual aircraft. Instead, the state space consists of the range of possibilities of the length of the queue at any given time, and the differential equations of the state dynamics govern how this queue grows or shrinks over time. There is no accounting, however, for which particular aircraft are present at any given time. Thus, the horizontal aggregation mechanism is not possible. The vertical aggregation mechanism is possible, and in a stochastic model each possible queue length is assigned some probability of prevailing at any particular time, so the vertically aggregated delay statistic generated represents the expected value of the delay incurred by aircraft during that time slice.

Fig. 1 shows an example of a cumulative demand curve (the upper curve, representing the number of flights that wanted to land by a particular time) and a cumulative supply curve (the lower curve, representing the actual number of flights that were allowed to land by a particular time, as constrained by the arrival capacity). The abscissa represents time, while the ordinate represents flight count, and the flights can be considered to be sorted in order of their desired arrival times.

During the time slice t_1 to t_2, flights labeled f_1 through f_2 landed. The total delay experienced by these flights over their lifetimes can be computed as the area of the horizontal band bounded by these two flight labels on the top and bottom, and by the two cumulative curves on the left and right. After dividing by the number of aircraft $f_2 - f_1$, the result is the average delay statistic that would have been reported in a database like ASPM. Because delays for individual flights are read from the figure as horizontal spacings between the two cumulative curves, we call this form of averaging delay "horizontal aggregation."

If one looks vertically at the same time slice, however, the band between the curves represents the total quantity of delay incurred by flights whose desired landing times occurred prior to the time slice in question, but whose actual landing times occur (or will occur) during or after that time slice. In this case, the total number of flights represented is $f_3 - f_1$ and the average delay statistic can be computed as the area of the vertical band divided by this number of flights. Again, this is the statistic traditionally (but not necessarily) drawn from

 ISBN: 978-0-615-20720-9

deterministic queuing models, and necessarily drawn from stochastic queuing models.

It should be clear from the figure that the two quantities can be quite different. Perhaps only a few flights are figured into both calculations, and even then the entirety of a flight's delay experience would be horizontally aggregated while only a portion of the delay would be captured with vertical aggregation during that time slice. Another way of thinking of the two methods is temporally: the horizontal aggregation method looks to the past, recording statistics about delays that have already occurred, while the vertical method looks to the present, by recording delays as they occur, but also to the future, because delays components yet to occur for those flights will be reported in later time slices. It is extremely important to note that both methods represent the "truth"; neither is more or less accurate than the other. The difference is simply in deciding which domain, in terms of time and flight identification, will be considered for aggregation and reporting during any particular time slice.

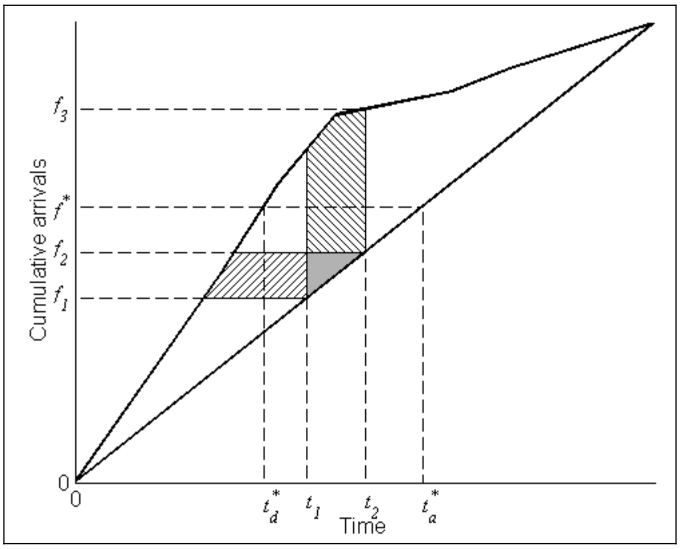

Figure 1. Vertical and horizontal integration of delay

When considering the specific flight f^* shown in Fig. 1, its desired landing time was t_d^* and its actual landing time was t_a^*. It would have contributed $t_2 - t_1$ units of delay to the vertical measure of aggregate average delay for that same time slice, and other amounts to other time slices. In the horizontal scheme, however, it would only have contributed to the measure recorded for the time slice containing time t_a^*, and the amount of delay contribution would have been $t_a^* - t_d^*$.

III. DELAY FILTERING

Because of the economic realities of the airline industry, individual aircraft are scheduled to operate several flights each day with little time between flights. Thus, if an aircraft suffers a delay early in the day, it becomes more likely that later flights operated by that same aircraft will also be delayed. When using real delay data to calibrate a queuing model, however, these propagated delays must be accounted for.

In this section, we describe an approach for identifying and removing these propagated delays from real delay data. The resulting statistics more clearly represent the queuing delay imposed on the aircraft. Additionally, we propose a technique to utilize this filtered data to produce a new "schedule" for each aircraft (and hence, for each airport). These schedules can be used as a better proxy for the true demand for resources as input to queuing models.

A. Procedure

The approach taken in this work and several others (see [2], [3]) has been to use individual flight records to trace aircraft by their tail numbers as they are routed from airport to airport over some period of time. These series of flights by a single aircraft are used to identify and remove propagated delay.

Thus, the first step in this process is to identify series of related flight data. Initially, data are grouped by tail number and sorted by departure day and time. However, because some flight records may be unavailable in the database, it may be infeasible to use all records for a single tail number as a single series. This phenomenon is evidenced by an arrival airport not matching the subsequent departure airport, indicating a missing flight record (e.g. caused by a "ferry" flight or data corruption). Data series are also considered broken if more than 24 hours elapse between an arrival and subsequent departures. From this procedure come several series of data for each tail number being examined.

Once these series of connected flights have been built, propagated delays must be distinguished from "new" delays. This process works by determining the best possible departure time for a flight, given the delay the previous flight experienced prior to its arrival. The best departure time is calculated as the maximum of two quantities: the scheduled departure time, or the previous (delayed) arrival time plus some minimum turn time, as shown in (2) and (3). The maximum of these two quantities is considered so as to prevent the best possible departure time from falling before the scheduled departure time. This would unfairly penalize flights relative to their schedule.

The minimum turn time is calculated in (1) as the minimum of the scheduled turn time and some parameter T_{turn}. In this analysis, only domestic flights were considered. Because these are generally operated by small or medium sized aircraft, the minimum turn time parameter T_{turn} was taken to be 40 minutes. An enhancement to be considered for future work using this delay filtering algorithm would be to consider variable minimum turn times, wherein the parameter might vary based upon the aircraft type, length of previous flight, airport in question, time of day, or some other factors.

Once the best departure time has been calculated, the best arrival time must be computed. In this work, the best arrival time was taken to be the sum of best departure time and the scheduled block time, as shown in (4). As mentioned previously, this block time does not represent the minimum time against which a queuing model might compare, as the scheduling carrier implicitly accounts for delay when scheduling the block time. However, estimating a true

 ISBN: 978-0-615-20720-9

minimum block time for a given flight may be a fairly complex endeavor, and as such, has been left for subsequent work.

Finally, the filtered arrival delay $D_{f,i}$ is computed as the maximum of zero and the difference between the best and actual arrival times. It is customary in aviation delay calculations to disregard negative delays, and we maintain that practice here, particularly because a queuing model would never predict negative delays.

This algorithm is stated below for a series of flights $i = 1, \ldots, I$, given the input data listed previous described.

Compute minimum turn time:

$$T_{turn,i} = \min\left[T_{turn}, \left(t_{sd,i} - t_{sa,i-1}\right)\right] \qquad \forall i \in \{2, \ldots, I\} \qquad (1)$$

Compute best departure time:

$$t_{bd,i} = t_{sd,i} \qquad i = 1 \qquad (2)$$

$$t_{bd,i} = \max\left[t_{sd,i}, \left(t_{aa,i-1} + T_{turn,i}\right)\right] \qquad \forall i \in \{2, \ldots, I\} \qquad (3)$$

Compute best arrival time:

$$t_{ba,i} = t_{bd,i} + T_{block,i} \qquad \forall i \in \{2, \ldots, I\} \qquad (4)$$

Compute filtered delay:

$$D_{f,i} = \max\left(0, t_{aa,i} - t_{ba,i}\right) \qquad \forall i \in \{2, \ldots, I\} \qquad (5)$$

Input data:
$t_{sd,i}$: Scheduled departure time for flight *i*
$t_{sa,i}$: Scheduled arrival time for flight *i*
$t_{aa,i}$: Actual arrival time for flight *i*
$T_{block,i}$: Scheduled block time for flight *i*

Once the filtered arrival delay $D_{f,i}$ has been computed for each flight, these records can be aggregated in either the horizontal or vertical methods previously mentioned. If they are aggregated horizontally by airport and time period, they will be comparable to those typically reported, but they will necessarily be lesser in magnitude. The case in which they are aggregated vertically will be discussed later.

If the aircraft counts are aggregated by best possible arrival time, it is possible to create a new "schedule" which better reflects the true demand for operations during that time period, from the perspective of the queuing model. For example, a queuing model being applied to ATL does not care if a flight had originally intended to arrive at ATL at 5 pm but due to delays two flight legs prior to that cannot even depart the airport immediately upstream of ATL until 5:30 pm. The real question for the queuing model is, given this penultimate status update, what would be the nominal arrival time for the aircraft at ATL. This data can be used as input for a delay prediction model to provide a better proxy for demand than the traditional schedule would.

B. Numerical Example

To illustrate the principles described above, a numerical example has been developed. The aircraft under consideration was routed as shown in Fig. 2. The detailed calculations for one flight leg are shown in Table I, while the scheduled and actual performance for its entire itinerary are shown in Table II.

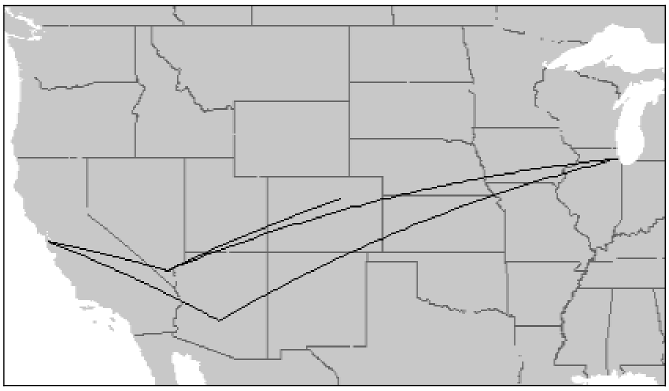

Figure 2. Example case routing

This aircraft suffers a fairly significant delay of 41 minutes on the second flight in this series. As a result, the best departure and arrival times computed for the subsequent flights are later than the scheduled ones. Further, the filtered flight delay must then be significantly less than the delay reported by traditional metrics.

To illustrate the steps of the algorithm described above, the calculations for the third flight leg (SFO – PHX) are presented here in Table I.

TABLE I. DELAY FILTERING COMPUTATIONS SFO – PHX FLIGHT

Compute minimum turn time:

$$T_{turn,i} = \min\left[40, \left(11:05 - 10:11\right)\right] = 40$$

Compute best departure time:

$$t_{bd,i} = \max\left[11:05, \left(10:52 + 40\right)\right] = 11:32$$

Compute best arrival time:

$$t_{ba,i} = 11:32 + 113 = 13:25$$

Compute filtered delay:

$$D_{f,i} = \max\left(0, 13:27 - 13:25\right) = 2$$

It is interesting to note that the schedule for each of these flights allowed for turns longer than the T_{turn} parameter of 40 minutes. As a result, each of the assumed turn times used to compute the best possible departure time was smaller than that which was scheduled. It should also be noted, however, that the average scheduled turn time was 48.5 minutes, while the average performed turn time was 44.3 minutes. Because of the time pressure of the delayed flight, the turns were performed faster than scheduled, and more closely matched the 40 minute parameter used in the algorithm.

 ISBN: 978-0-615-20720-9

TABLE II. DELAY FILTERING EXAMPLE DATA

From	To	Schedule			Actual			Best			Delay	
		Turn	Dep.	Arr.	Turn	Dep.	Arr.	Turn	Dep.	Arr.	Reported	Filtered
DEN	LAS	-	7:10	7:55	-	7:05	7:56	-	7:10	7:55	1	1
LAS	SFO	0:45	8:40	10:11	0:39	8:35	10:52	0:40	8:40	10:11	41	41
SFO	PHX	0:54	11:05	12:58	0:37	11:29	13:27	0:40	11:32	13:25	29	2
PHX	ORD	0:50	13:48	19:14	0:55	14:22	19:36	0:40	14:07	19:33	22	3
ORD	LAS	0:45	19:59	21:49	0:46	20:22	22:10	0:40	20:16	22:06	21	4

IV. VERTICAL INTEGRATION

There are many ways in which individual flight delays can be aggregated. The most familiar category of metrics involves summing delays across flights arriving at a given airport (or set of airports) during a particular time period. However, one must be precise when describing exactly which data are summed for the given time period.

In the traditional metrics reported in the ASPM and other systems, delays are grouped according to the time at which flights arrived. Regardless of when those delays were accrued, they are assigned to the period of arrival under consideration. The essence of the vertical integration technique, however, is to sum delays that are accrued during a given time period, regardless of when the affected flights arrive.

A. Procedure

The first part of this procedure is to establish at what time delay begins accruing on a flight. Establishing this baseline allows the delay to be assigned to bins beginning at that time. This assumption must be carefully examined, lest delay be assigned to the incorrect time bins. In this work, we assume that delay begins accruing when the nominal, or best possible, arrival time has passed, and the aircraft has not yet arrived at its destination. This best possible arrival could be calculated in many ways, depending upon the assumptions about departure and flight times that were applied. Based upon the delay filtering analysis presented previously, we will use the best possible arrival time calculated as part of that algorithm.

The first step in calculating these delays is to divide each day into a series of time bins, each bounded by some numbers t_p and t_{p+1}. Let L and U define the upper and lower bounds for the delay accrual period. In this case, these bounds are the best arrival time and the actual arrival time, respectively. Then, find the first bin l into which the flight i contributes delay, as shown in (6).

$$l = \max\left\{ p \mid L - t_p \geq 0 \right\} \tag{6}$$

Next, find the last bin u into which the flight i contributes delay, as shown in (7).

$$u = \max\left\{ p \mid U - t_p \geq 0 \right\} \tag{7}$$

Then, for each bin $p \in \left\{ l, \ldots, u \right\}$, apply the following four logical tests to determine the delay accrual $D_{p,i}$ from flight i into bin p.

(a.) IF $L \geq t_p \;\&\; L < t_{p+1} \;\&\; U < t_{p+1} \;\&\; L < U$

THEN $D_{p,i} = U - L$

(b.) ELSE IF $L \geq t_p \;\&\; L < t_{p+1} \;\&\; U \geq t_{p+1}$

THEN $D_{p,i} = t_{p+1} - L$

(c.) ELSE IF $U \geq t_p \;\&\; U < t_{p+1} \;\&\; L < t_p \;\&\; L < U$

THEN $D_{p,i} = U - t_p$

(d.) ELSE IF $L < t_p \;\&\; U \geq t_{p+1}$

THEN $D_{p,i} = \left(t_{p+1} - t_p \right)$

The $L<U$ condition is applied to exclude those cases in which the flight arrives before its best possible arrival time. In those cases, the new calculated delay would be negative. We treat these cases as having accrued zero delay.

Fig. 3 illustrates each of these logical tests, and the specific case of L and U that they approach. The hatched area in the figure shows the delay accrual period for the flight. Case (a.) applies when both L and U fall in the same time bin. Case (b.) applies when L is in the current bin, but U is in any later one. Case (c.) applies when L is in a previous time bin, but U is in the current one. Case (d.) applies when L is in an earlier time bin, and U is in a later one.

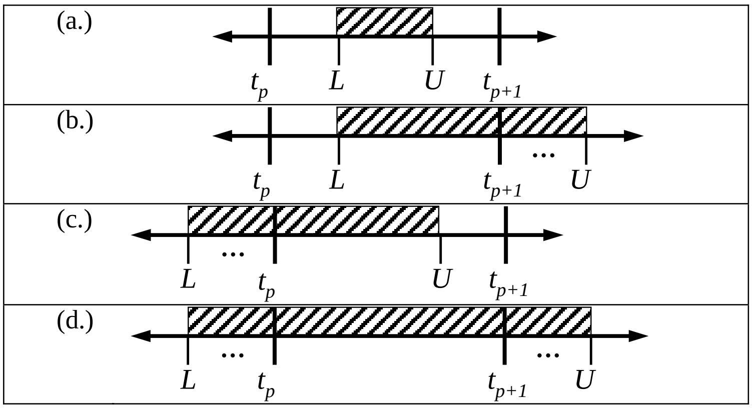

Figure 3. Logical cases for binning delays

B. Numerical Example

This algorithm is illustrated here by examining a fictitious set of flights shown in Table III, and represented graphically in Fig.. 4. Assume that the delay filtering algorithm previously described has been applied to a larger dataset, and that these flights destined for ORD were extracted. The best departure and arrival times, as well as the actual arrival times, are shown.

 ISBN: 978-0-615-20720-9

The filtered delay is calculated as the difference between the best possible and actual arrival times.

TABLE III. VERTICAL INTEGRATION EXAMPLE DATA

From	To	Best		Actual Arrival	Filtered Delay
		Departure	*Arrival*		
ATL	ORD	12:19	1:23	1:38	15
BWI	ORD	12:11	1:26	1:32	6
LGA	ORD	12:05	1:29	1:41	12
CLT	ORD	12:03	1:00	1:21	21
OKC	ORD	12:00	1:15	1:30	15

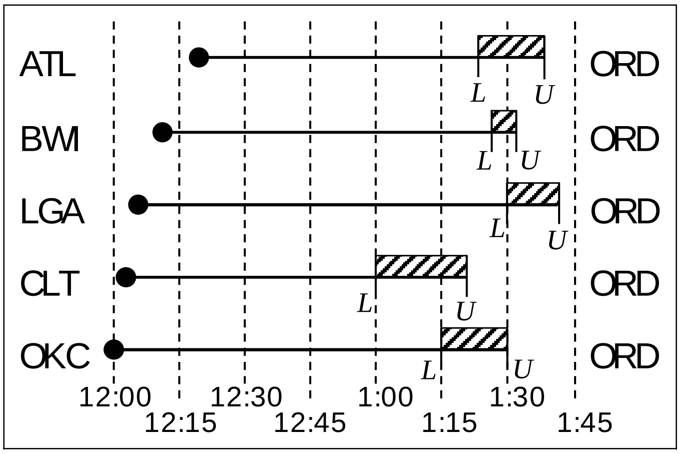

Figure 4. Vertical integration example data

As an example, apply the various tests on the first flight shown above, that from ATL to ORD. The first bin l to consider is the 1:15 bin, and the last bin u is the 1:30 bin. Each of the logical tests is evaluated for this flight and the results shown in Table IV.

TABLE IV. VERTICAL INTEGRATION COMPUTATIONS FOR ATL – ORD FLIGHT

Test	1:15 bin		1:30 bin	
	Test Result	$D_{p,i}$	Test Result	$D_{p,i}$
(a.)	FALSE	-	FALSE	-
(b.)	TRUE	7	FALSE	-
(c.)	FALSE	-	TRUE	8
(d.)	FALSE	-	FALSE	-

Upon evaluating the logical tests for each of these flights, the data are summed across time bins, and the results summarized in Table V. As expected, the reported delay differs significantly from the delay actually accrued by all flights in each period.

TABLE V. VERTICAL INTEGRATION EXAMPLE SUMMARY STATISTICS

Time period		Arrival Count		Total Arrival Delay	
Begin	End	Sch.	Actual	Reported	Accrued
1:00	1:14	1	0	0	15
1:15	1:29	3	1	21	31
1:30	1:44	1	4	48	22

V. RESULTS

The delay filtering and vertical integration algorithms were incorporated and applied to a large test dataset to provide comparison data for calibrating a queuing model. Several airports were considered.

A. Delay Models

For this work, we used a stochastic queuing model with non-homogeneous Poisson arrivals and Erlang services times ($M(t)/E_k(t)/1$). This model is solved analytically using the DELAYS software developed at MIT (see [4], [5]). It is specifically designed to estimate the delay incurred by aircraft on landing at an individual airport given the capacity and demand profiles over specific time periods. Previous work has focused considerable attention on the accuracy of the approximation scheme used by this software. This work will demonstrate that DELAYS provides a suitable queuing model for airport arrival operations.

B. Input Data

The individual flight records were obtained from the Bureau of Transportation Statistics (BTS). The BTS database includes records for certified US air carriers that account for at least one percent of domestic scheduled passenger revenues. Other sources of individual flight records could be used as well, but the BTS data provides excellent coverage of operations at most of the largest airports in the US.

The Airport Arrival Rates used as capacities for the DELAYS model were drawn from the ASPM system. The demands used as input for the DELAYS model were not the scheduled demand, but rather were summed using the best possible arrival time as the scheduled arrival time.

C. Case Studies

Data from numerous airports were examined in this work. The results for several are shown here, but others are available from the authors.

The "Reported" category refers to the horizontally aggregated data typically reported. The "Filtered" category shows the results of filtering out propagated delays, but aggregating in the traditional horizontal manner. The "Filt/Vert" category shows the results of both filtering the data, and aggregating it by the period in which it was accrued.

ISBN: 978-0-615-20720-9

1) Atlanta Hartsfield-Jackson International Airport (ATL)

ATL is a very large and busy airport serving as the hub for several carriers. Demand is frequently at or near capacity.

Fig. 5 shows results for a sample month during 2004. The first thing to note is that the reported delays are almost always higher than all other metrics. The filtered delays fall slightly below the reported delay, but follow the same series of peaks and valleys. Particularly at the end of the day, the gap between these two is large, as should be expected. The vertically integrated and filtered data suggest an amount of total delay similar to the filtered delay, but have peaks and valleys that more closely follow those of the DELAYS series, which are nearly almost lower.

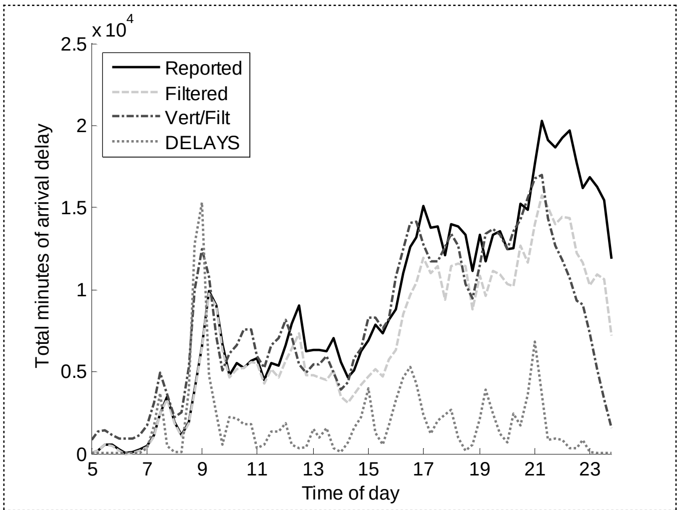

Figure 5. ATL: February. 2004

Fig. 6 shows three sets of pairwise correlations, between the three delay quantities mentioned above, and the delays predicted by the DELAYS queuing model over each month in 2004. The correlations between the vertically integrated and filtered data and the DELAYS output (the gray bars) were uniformly higher than those for all other metrics. The figure only shows results for ATL, but this conclusion held true for every airport examined.

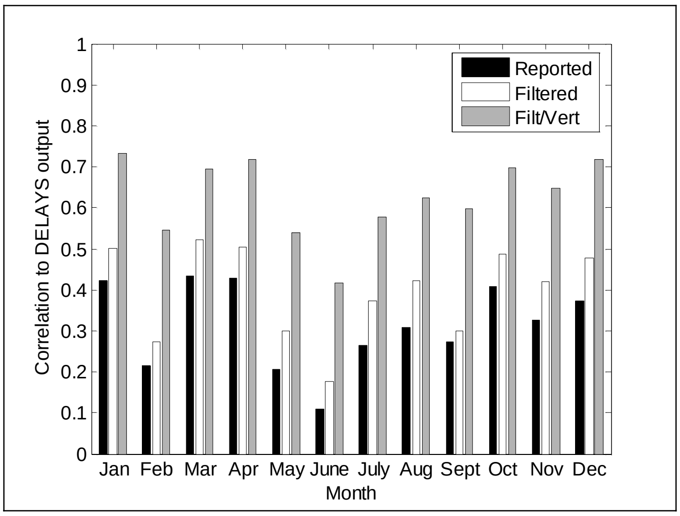

Figure 6. ATL 2004: Monthly Correlations

2) Detroit Metropolitan Wayne County Airport (DTW)

The results for DTW were fairly similar to those for ATL. An interesting feature of the DTW results, which was present to a much lesser degree for ATL, is the correspondence of peaks and valleys in the monthly data shown in Fig. 7. The peaks for the reported and filtered series correspond quite well, as should be expected, because they are both horizontally aggregated. In addition, the peaks for the vertically integrated filtered data and the DELAYS model correspond quite well. The interesting feature here, however, is that the peaks for the first pair of data lag those for the second pair. This exhibits the exact feature espoused earlier in the paper, which is that the delay model will show delay as it is accumulated, while the reported statistics will show it as the aircraft arrive.

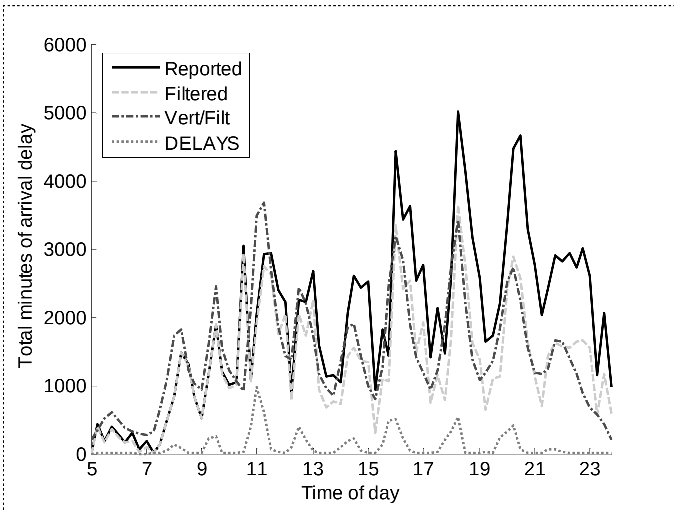

Figure 7. DTW: December, 2004

As was the case for ATL, the correlations between the vertically integrated filtered data and the DELAYS outputs are uniformly and significantly higher than those of any other metric. This suggests that the proposed methods provide data that corresponds better with the DELAYS model output, and we would expect this same conclusion to hold for other queuing models.

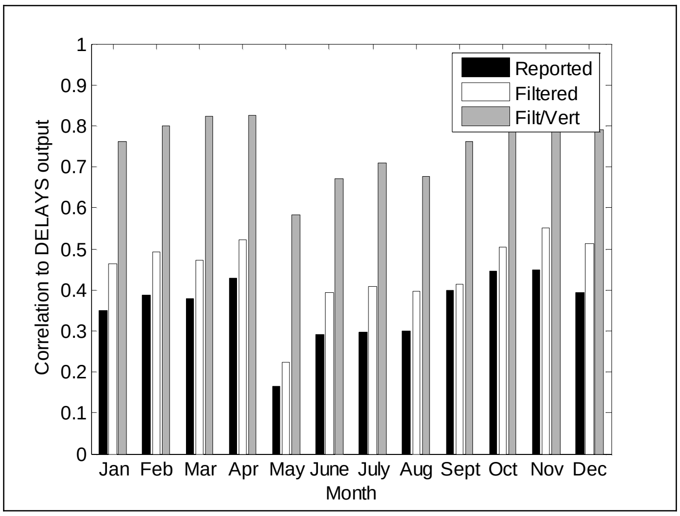

Figure 8. DTW 2004: Monthly correlations

ISBN: 978-0-615-20720-9

VI. CONCLUSIONS

Several schemes were presented in this paper to help understand the relationship between operational and queuing model data in aviation systems. In their native formats, the data have slightly different contextual meanings, and this makes direct comparison troublesome. They can be rectified, however, by the methods discussed in the paper.

The first method discussed for bringing the data sources into agreement was the application of filtering techniques. These are useful in removing the effects of delay propagated between flights using the same aircraft. This technique removed some portion of this delay, and produced data that showed a stronger correlation with predicted results.

The second technique shown in this paper was a different scheme of aggregation than is typically used for aviation delay data. The methodology proposed allows for delays to be reported in the time bin in which they are accrued, rather than the time bin in which the flight arrives. This technique, combined with the first, produced results that show a very strong correlation to the predicted delays.

These two techniques have myriad applications in aviation system planning and modeling. Both are very useful in calibrating and understanding delay prediction models. In addition, they encourage the reader to consider the nature of the delay reporting mechanisms in use.

REFERENCES

[1] Newell, G.F. 1982. Applications of queuing theory, 2nd Edition. Chapman and Hall, New York, NY.

[2] Lovell, D.J., A.M. Churchill, A.R. Odoni, A. Mukherjee, Ball, M. O., 2007, "Calibrating Aggregate Models of Flight Delays and Cancellation Probabilities at Individual Airports," *Proceedings of 7th USA/Europe Air Traffic Management R&D Seminar*, Barcelona, Spain.

[3] Mukherjee, A., D.J. Lovell, M.O. Ball, A.M. Churchill, A.R. Odoni. 2007. Estimating long-term average flight delays and cancellation rates for an airport, unpublished.

[4] Kivestu, P. 1976, "Alternative Methods of Investigating the Time-dependent M/G/K Queue," S.M. Thesis, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA.

[5] Malone, K., 1995, "Dynamic Queuing Systems: Behavior and Approximations for Individual Queues and Networks," Ph.D. Thesis, Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA.

ISBN: 978-0-615-20720-9